The Journal of Open Source Software

The sspm R package: spatial surplus production models for the management of northern shrimp fisheries

Valentin Lucet [*1] and **Eric J. Pedersen** [†1,2]

1 Department of Biology, Concordia University, Montreal, CA **2** Department of Biology, Memorial University of Newfoundland, St. John's, CA

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Statement of need

1. Population models, in particular fisheries productivity models, rarely integrate important spatially-structured ecosystem drivers
2. The Northern Shrimp stock in the Newfoundland and Labrador Shelves currently lacks a population model
3. Current SPM models are rarely spatially explicit and usually cannot account for relevant ecosystem drivers
4. Fisheries managers lack user-friendly, flexible tools to implement and apply SSPMs

Summary

Population modelling is an exercise of interest within environmental sciences and adjacent fields. Early population models such as the logistic model assumed that while the abundance of a population might change over time, the environmental conditions governing parameters affecting that rate of change (such as the maximum rate of growth or the carrying capacity of the population) stay constant over time (Gotelli, 2008). Modern population models increasingly acknowledge the non-stationary nature of wild populations, and work to incorporate environmental fluctuations into dynamic models (Thorson et al., 2015, 2017). Population models designed to answer applied resource management questions, such as fisheries stock models, increasingly address how the dynamics of stocks vary across space and time.

Resource managers are becoming increasingly interested in how variation in ecosystem factors such as predator abundance and abiotic variables impact the spatiotemporal variability of population parameters such as productivity (Szuwalski & Hollowed, 2016; Zhang et al., 2021). Further, treating spatially structured stocks as single unstructured stocks can lead to substantially biased estimates of population change (Thorson et al., 2015). However, stock models that explicitly incorporate spatial dynamics and time-varying ecosystem variables still rare in fisheries science, despite the push for more ecosystem-based management methods in fisheries management (Berkes, 2012; Crowder et al., 2008; Tam et al., 2017).

Surplus production models (SPMs) are one of the classic models used in fisheries, and are based on modelling changes in the total biomass of a stock in a given location over time (i.e. *surplus production*) as a function of current stock abundance and fishing pressure (Walters et al., 2008). They are useful in data-poor contexts where the age and sex structure of the population is not accessible, or when age- or length-structure do not change substantially over time (Prager, 1994; Punt, 2003). Classically, SPMs assumed single unstructured stocks with

*co-first author
†co-first author

purely logistic dynamics (Walters et al., 2008), and as such have been viewed as a limited tool for modelling more complex stocks.

Two limitations of classic SPMs is the assumption of spatially constant productivity, and that stock productivity is affected only by stock abundance and fishing. These assumptions ignore the effect of global changes that are affecting fisheries, such as climate change, that affect the growth rates of stocks independently of fishing pressure. One example is that of the Northern Shrimp (*Pandalus borealis*) in the Newfoundland and Labrador Shelves, which has undergone several periods of large-scale biomass change in the last two decades, despite a relatively constant harvest regime (DFO, 2019). These stocks currently lack a population model to understand the drivers of this change, and to predict how fishing pressure and changing environmental conditions may affect future abundance in the region. In the context of climate change and shifting ranges, fisheries productivity is likely to be a moving target (Karp et al., 2019), and managers need better methods that account for varying productivity (Szuwalski & Hollowed, 2016).

Population models like SPMs usually fall under two categories: process-based and statistical. Process-based models often rely on differential or difference equations and are based on replicating the underlying processes (e.g., predation, recruitment, dispersal) driving population dynamics. Statistical models instead fit a regression model to time series of population abundances, abundance indices, or productivities, with some assumed error distribution for variation around predictions. This allows for estimation of parameter uncertainty and ranges of model predictions, and for flexibly incorporating potential ecosystem drivers into models (Plagányi et al., 2014). Statistical models also allow for straight-forward estimation of spatial variation in population parameters such as maximum productivity or density dependence from data, in the absence of theory predicting how these parameters should vary. In this paper, we use a statistical approach to fitting SPMs using Generalized Additive Models (GAMS), estimated using the `mgcv` R package (Wood, 2017) as the backend. We apply this approach to the population of Northern Shrimp of the Newfoundland and Labrador Shelves, leveraging the smoothing properties of GAMs to account for varying productivity across time and space. The resulting model is a spatial SPM (SSPM), implemented via a R package `sspm`.

While the initial application of this model was modelling Newfoundland and Labrador Northern Shrimp stocks (Pedersen et al., 2022), The R package `sspm` is designed to make spatially-explicit surplus production models (SSPM) simpler to estimate and apply to any spatially structured stock. The package uses GAMs to estimate spatiotemporally varying biomass, and to estimate SSPMs based on changes in fitted biomass, observed catch, and spatially structured environmental predictors. It includes a range of features to manage biomass and harvest data. Those features are organized in a stepwise workflow, whose implementation is described in more detail in Figure 1.

1. Discretization and aggregation of spatially structured observations into discrete patches, with a range of methods of discretization (random or custom sampling, Voronoi tessellation or Delaunay triangulation).
2. Spatiotemporal smoothing of biomass and environmental predictors using GAMs.
3. Computation of surplus productivity based on biomass density and fishing effort.
4. Fitting of SSPMs to productivity data with GAMs.
5. Visualization of results, including confidence and prediction intervals.
6. One-step-ahead prediction of biomass for model validation and scenario-based forecasting.

Although it was developed in a fisheries context, the package is suitable to model spatially-structured population dynamics in general.

Package design

The package follows an object oriented design, making use of the S4 class systems. The different classes in the package work together to produce a stepwise workflow (Figure 1).

The key workflow steps are:

1. Delineation of the boundary of the region of interest for the model. Boundary data is provided as a shapefile and converted into a `sspm_boundary` object with a call to `spm_as_boundary()`.
2. The region within the boundary is discretized into patches with the `spm_discretize()` function, creating a `sspm_discrete_boundary` object.
3. The `spm_as_dataset()` function turns user-provided data frames of raw observations into `sspm_dataset` objects that explicitly track locations, data types, and aggregation scales for each input. `sspm` recognizes three types of data: **trawl** (i.e. biomass estimates from scientific surveys), **predictors**, and **catch** (i.e. harvest).
4. The `spm_smooth` function use spatiotemporal GAM models to smooth the biomass and predictor data, based on the spatial structure from `sspm_discrete_boundary`. The user specifies a GAM formula with custom smooth terms. The output is another `sspm_dataset` object with a `smoothed_data` slot which contains the smoothed predictions for all patches.
5. The `spm_aggregate_catch` function aggregates catch into patches and years, and calculates patch-specific productivity for each year as the ratio of estimated biomass density plus catch from the next year, divided by estimated biomass density of the current year. The result is returned as a `sspm_dataset`.
6. The `sspm` function combines productivity and predictor datasets into a single dataset. Additionally, the user may create lagged versions of predictors with `spm_lag()` and split data into testing and training sets for model validation with `spm_split()` at this stage.
7. The `spm()` function is used to fit a SSPM model to the output of step 6, using a GAM model with custom syntax able to model a range of SSPMs. The output is an `sspm` object.
8. Predictions from the fitted model can be obtained using the built-in `predict()` method, and plots with the `plot()` method.

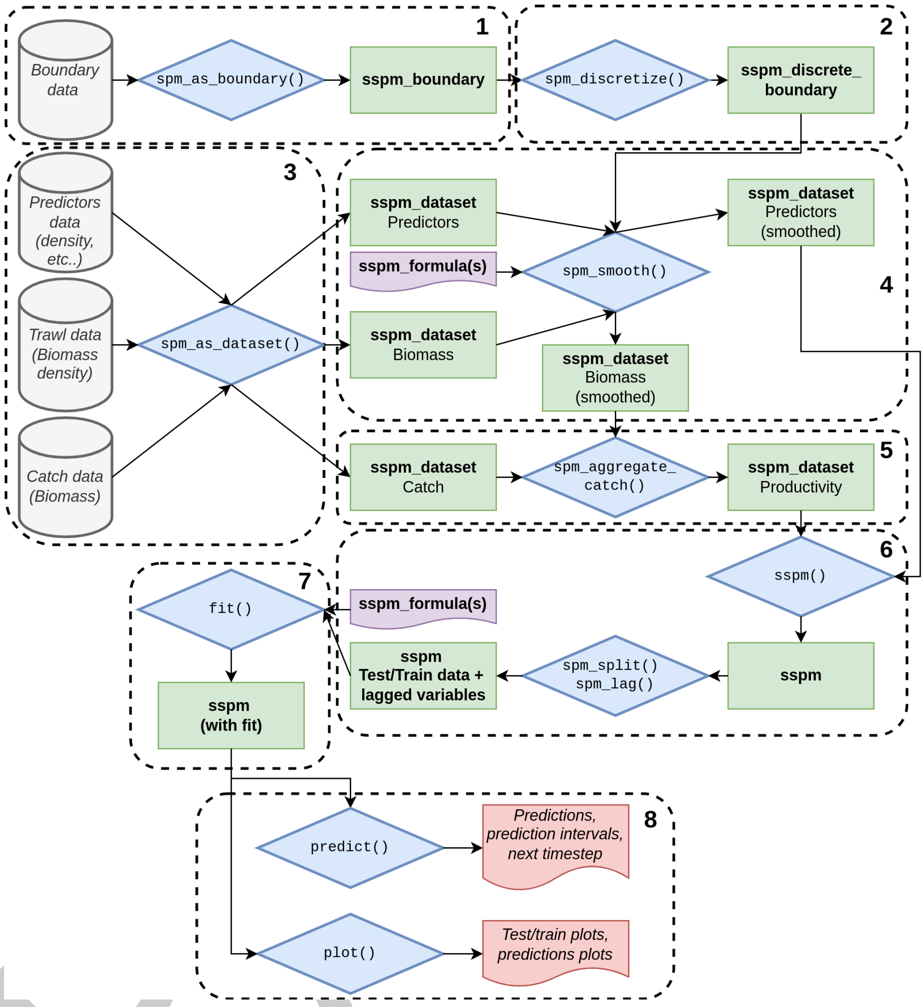

Figure 1: The sspm workflow.

Connections to other spatiotemporal stock assessment approaches

The general approach used by `sspm` of using a statistical model to estimate spatiotemporally varying population varying biomass indices is closely connected to approaches used by other modern model-based spatial abundance estimation software, such as the `VAST` R package (Thorson, 2019) and the `sdmTMB` R package (Anderson et al., 2022). Our method shares the same approach as both `VAST` and `sdmTMB` of using spatially explicit models to estimate local biomass density (Figure 1 steps 1,2, and 4), then aggregating up from those models to predict aggregate stock-level metrics such as total biomass and productivity (Figure 1 steps 8). The multiplicative surplus production model used by `sspm` is also similar to the vector-autoregessive model for biomass changes used by these two packages, as both `VAST` and `sdmTMB` can model local temporal changes as autoregressive processes on the link-scale of a generalized linear model (Thorson, 2019).

One major difference between the `sspm` package and other model-based spatiotemporal modelling packages is its special-purpose nature. The `spm_smooth` function uses a computationally simpler (although somewhat less flexible) Conditional Autoregressive (CAR) model (Rue & Held, 2005) for modelling spatial variation in covariates and biomass, as compared to the more complex spatial random effects possible with `VAST` and `sdmTMB`; this has the advantage of computational speed and less user knowledge of how to set up complex spatial grids, although it is less flexible. This means that `sspm` should be easier to adapt to novel fisheries than more complex packages that require more user modelling knowledge.

The other benefit of `sspm`, relative to other modelling packages, is the ability to model productivity rates directly (Figure 1 steps 5 and 6), rather than implicitly via an auto-regressive processes as used in `VAST` or `sdmTMB`. This means that it is possible in `sspm` to model nonlinear relationships between environmental covariates and productivity, or to easily include factors such as time-lagged effects of predictors on productivity in a given year. This approach does, however, sacrifice the ability to propagate measurement error into uncertainty about rates of change. One of the future directions for development of this package is to include variance propagation methods into the surplus production modelling step.

Acknowledgements

This research was supported by the Canadian Department of Fisheries of Oceans (DFO) Sustainable fisheries Science Fund, and by a Discovery Grant from the Canadian Natural Sciences and Engineering Research Council (NSERC) to E. J.Pedersen. We thank Fonya Irvine and John-Philip Williams for their help in testing the package and providing feedback on model implementation.

References

Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). *sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields* (p. 2022.03.24.485545). bioRxiv. https://doi.org/10.1101/2022.03.24.485545

Berkes, F. (2012). Implementing ecosystem-based management: Evolution or revolution? *Fish and Fisheries*, *13*(4), 465–476. https://doi.org/10.1111/j.1467-2979.2011.00452.x

Crowder, L. B., Hazen, E. L., Avissar, N., Bjorkland, R., Latanich, C., & Ogburn, M. B. (2008). The Impacts of Fisheries on Marine Ecosystems and the Transition to Ecosystem-Based Management. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 259–278. https://doi.org/10.1146/annurev.ecolsys.39.110707.173406

DFO. (2019). *An assessment of Northern Shrimp (*Pandalus borealis*) in Shrimp Fishing Areas 4–6 and of Striped Shrimp (*Pandalus montagui*) in shrimp fishing area 4 in 2018* (Technical Report No. 2019/027; p. 23). Canadian Science Advisory Secretariat (CSAS).

Gotelli, N. J. (2008). *A Primer of Ecology* (Fourth Edition). Oxford University Press. ISBN: 978-0-87893-318-1

Karp, M. A., Peterson, J. O., Lynch, P. D., Griffis, R. B., Adams, C. F., Arnold, W. S., Barnett, L. A. K., deReynier, Y., DiCosimo, J., Fenske, K. H., Gaichas, S. K., Hollowed, A., Holsman, K., Karnauskas, M., Kobayashi, D., Leising, A., Manderson, J. P., McClure, M., Morrison, W. E., … Link, J. S. (2019). Accounting for shifting distributions and changing productivity in the development of scientific advice for fishery management. *ICES Journal of Marine Science*, fsz048. https://doi.org/10.1093/icesjms/fsz048

Pedersen, E. J., Skanes, K., Le Corre, N., Koen-Alonso, M., & Baker, K. (2022). *A new spatial ecosystem-based surplus production model for SFA 4-6 Northern Shrimp* (Can. {Sci}. {Advis}. {Sec}. {Res}. {Doc}. No. 2022/nn; pp. in press).

Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D. M., Smith, F., Bayliss, P., Haywood, M., Lyne, V., & Rothlisberg, P. C. (2014). Multispecies fisheries management and conservation:

Tactical applications using models of intermediate complexity. *Fish and Fisheries*, *15*(1), 1–22. https://doi.org/10.1111/j.1467-2979.2012.00488.x

Prager, M. H. (1994). *A suite of extensions to a nonequilibrium surplus-production*. 17.

Punt, A. E. (2003). Extending production models to include process error in the population dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*(10), 1217–1228. https://doi.org/10.1139/f03-105

Rue, H., & Held, L. (2005). *Gaussian Markov random fields: Theory and applications*. CRC press.

Szuwalski, C. S., & Hollowed, A. B. (2016). Climate change and non-stationary population processes in fisheries management. *ICES Journal of Marine Science*, *73*(5), 1297–1305. https://doi.org/10.1093/icesjms/fsv229

Tam, J. C., Link, J. S., Rossberg, A. G., Rogers, S. I., Levin, P. S., Rochet, M.-J., Bundy, A., Belgrano, A., Libralato, S., Tomczak, M., Wolfshaar, K. van de, Pranovi, F., Gorokhova, E., Large, S. I., Niquil, N., Greenstreet, S. P. R., Druon, J.-N., Lesutiene, J., Johansen, M., … Rindorf, A. (2017). Towards ecosystem-based management: Identifying operational food-web indicators for marine ecosystems. *ICES Journal of Marine Science*, *74*(7), 2040–2052. https://doi.org/10.1093/icesjms/fsw230

Thorson, J. T. (2019). Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, *210*, 143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Thorson, J. T., Jannot, J., & Somers, K. (2017). Using spatio-temporal models of population growth and movement to monitor overlap between human impacts and fish populations. *Journal of Applied Ecology*, *54*(2), 577–587. https://doi.org/10.1111/1365-2664.12664

Thorson, J. T., Skaug, H. J., Kristensen, K., Shelton, A. O., Ward, E. J., Harms, J. H., & Benante, J. A. (2015). The importance of spatial models for estimating the strength of density dependence. *Ecology*, *96*(5), 1202–1212. https://doi.org/10.1890/14-0739.1

Walters, C. J., Hilborn, R., & Christensen, V. (2008). Surplus production dynamics in declining and recovering fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(11), 2536–2551. https://doi.org/10.1139/F08-170

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, 2nd Edition* (2nd ed.). CRC Press. ISBN: 978-1-4987-2837-9

Zhang, F., Reid, K. B., & Nudds, T. D. (2021). The longer the better? Trade-offs in fisheries stock assessment in dynamic ecosystems. *Fish and Fisheries*, *22*(4), 789–797. https://doi.org/10.1111/faf.12550